Quality Control: The Great Myth

Thomas L. Francoeur
Atlantic Ecotechnologies/TEG Northeast
160 Longwoods Road; Cumberland, ME 04021
Phone: (207) 829-6752 Fax: (207) 829-6754

ABSTRACT

Both data generators and data users are under economic pressures to drive down the cost of their respective services. This pressure forces data generators take shortcuts, and data users circumvent the Data Quality Objective (DQO) process. This combination of factors is very dangerous and has lead to an untold number of situations where the end user's absolute confidence in environmental data is unwarranted. Confidence in environmental data is rationalized through laboratory certification and the mere performance of Quality Control Procedures as an assurance of data quality rather than a measure of data quality. Ironically, because these short cuts can so dramatically impact price, both generators and users are rewarded by receiving additional work. This vicious cycle has lead to a proliferation of "data time bombs" where data go on to be used in reports for what may be an inappropriate use. This paper will discuss the basics of the DQO process and how data should be deemed usable for a given use. This paper will further discuss how the DQO process does not have to be a cumbersome and complex process, but rather an essential component of an environmental investigation, and will illustrate the potential negative result by discussing several examples of "data time bombs".

INTRODUCTION

The inappropriate use (or non-use) of quality control procedures has lead to a false sense of security with respect to the "quality" of analytical data. Often there is a disconnect between QA/QC, Data Quality Objectives (DQOs), and the ultimate verification and validation of data quality. Environmental data often go on to their final use without verification or validation. Moreover, when one of these processes does occur, it is wrongly viewed as an assurance of quality data, rather than as nothing more than an instrument to measure and evaluate data quality and data usability. When credible data review does not occur, the potential exists for incorrect conclusions to be drawn from analytical data. These data, and their subsequent conclusions, often exist in reports for years until the "data time bomb" explodes in a fury of backpedaling, accusation, expense, and reinvestigation.

In an era of intense laboratory (both fixed base and field) competition, automation and QA/QC have been viewed as a substitute for experienced personnel as a means of driving down unit costs. Although QA/QC systems are essential to monitor analytical processes, they are not substitutes for competent personnel because there are many actions which can impact data quality which do not have a germane QC procedure. This problem is exacerbated by data users who are not interested in how data were generated but only in the numbers themselves, and as a result, circumvent the data review process. Dangerous assumptions are being made as to the "quality" of analytical results. It is assumed that if QC procedures were used, or a laboratory is "certified", than the data are necessarily of high quality and can be accepted without question. Conversely, environmental data generated in the field are often further stigmatized by the label "screening" when in fact they may be of objectively higher quality (i.e., more usable) than their fixed laboratory counterpart. This assumption that fixed-base laboratory data are "good" and field data are "screening" (i.e., of limited use or suspect)

exists even though little examination of QC procedures by the data user occurs in either case.

The use of mobile laboratories is a tremendous opportunity to eliminate the disconnect between data generators and data users. By performing analyses in the field, the chemist understands how the data are to be used and the project manager understands how the data were generated. Performance based procedures can be established for a given data use, and necessary and relevant QC procedures can be monitored by both the laboratory and the data user so corrective actions (or perhaps even optimization actions) can be implemented in time for them to make a difference. Data usability does not end with laboratory QA/QC or even certification. Rather, procedures have to be defined for a given site, quality control tests need to be implemented to monitor these procedures, skilled personnel need to be present to execute them, and finally, the data need to be assessed through some independent means.

HISTORY AND DEFINITIONS

A tremendous amount of confusion surrounds the definitions of and the relationships between terms such as data quality objectives, data verification, data validation, and data defensibility. To better understand what goes into generating usable data, the following terms are defined presented in the order they should occur in the chronology of a project.

Data Quality Objectives (DQOs)

Although well documented[1,2,3,4,5], the DQO process is not widely understood nor widely implemented. Simply, DQOs specify requirements for analytical data that are clear and unambiguous concerning the intent of an investigation and the data parameters necessary to achieve that intent. These objectives are stated in both qualitative terms concerning the end use of the data, as well as quantitative terms with respect to precision, accuracy, representativeness, comparability and completeness (PARCC). Unfortunately, these terms are often inserted into work plans to satisfy the requirements of various regulatory agencies. After the data generation process, however, the subsequent values for the individual parameters are never calculated.

DQOs insure that the proper methods and procedures (including method modifications) are in place to insure project success with respect to detection limits, practical quantitation limits, Applicable or Relevant and Appropriate Requirements (ARARs), action limits, compound specificity, compound selectivity, reproducibility, false positives, and false negatives which are all incorporated into the aforementioned PARCC parameters. DQOs insure that procedures are implemented to monitor these parameters, and that the process will ultimately generate useable data.

Data Verification

Simply, data verification occurs after the data analyses are completed. Data verification is a rigorous process whereby quality control parameters are evaluated against a set of predetermined criteria or functional guidelines. The scope of data verification is limited to evaluation of an analytical procedure against its internal Quality Control procedures, such as a standard operating procedure (SOP) or standard method. It will also track and confirm the paper trail of a given sample. Data verification often involves the examination of such quality control parameters as holding times, surrogates, calibrations, method blanks, field blanks, duplicates, matrix spikes, matrix spike duplicates, and laboratory control samples[6,7]. Data are qualified as estimated or rejected based on the performance of the laboratory during the receipt and analysis of environmental samples, but there is very little flexibility to alter the results of data verification based on professional judgment or experience. It should be noted that the data verification process states nothing qualitatively with respect to the usability of the data. Data use is not taken into account in the data verification process.

Data Validation

Data validation is the final step in a process whereby data usability is assessed. Unfortunately, data validation is often misunderstood and confused with data verification, and as a result does not often actually occur. It is important to understand the validation of data does not guarantee valid data. The reviewer responsible for data validation must be knowledgeable of the site in question and must have a handle on a variety of disciplines such as fate and transport, hydrogeology, statistics, and project management, as well as analytical chemistry. Based on this larger "world view", the data reviewer will assess the usability of the data. Based on these larger inputs, the data reviewer may do such extreme things as reject data previously unflagged as a result of verification, or conversely, accept data as usable which were flagged as rejected during the data verification process. This can occur because the reviewer understands how the data were generated as well as how the data need to be used.

Defensible Data and Usable Data

In part, data defensibility is a conclusion based on data validation. There is a major misconception between both buyers and sellers of environmental services as to what constitutes "defensible" data. Data cannot be deemed "defensible" until a data use has been defined, and the data are consequently either appropriate (defensible) or inappropriate for that use. Too often "defensible" is characterized as "usable under any circumstances", when in fact the EPA characterizes defensible as "the ability to withstand any reasonable challenge related to the veracity, integrity, or quality of the logical, technical, or scientific approach taken in a decision making process"[8]; therefore a determination of defensibility (i.e., usability) cannot be made in the absence of a decision making process (i.e., a data use). There exists a common misconception that data must be generated in a "CLP-like" fashion in order to stand up in court, when in fact even hand-held meter readings are defensible provided they meet the need, and requirements, of a predefined data use.

LABORATORY CERTIFICATION

Simply, certification is not a substitute for the verification and validation process. Unfortunately, certification has become a substitute for data quality because it takes much less effort and expertise to determine if a laboratory is certified than to undertake the DQO process. The requirements for certification vary greatly from state to state, but in general, when a laboratory is "certified", it met the minimum requirements to perform a specific operation. Certification is not a blanket approval of the laboratory, its equipment and its personnel; nor does certification guarantee valid, let alone usable, data. A certification reflects a snapshot in time, rather than a reflection of the more general processes conducted by a laboratory. Most importantly, however, there is no consideration on the data generator's part of ultimate data use. This is why an independent data validation process is so critical. Only after validation will you have data of *known* quality (not necessarily good quality). Rather, certification indicates that a laboratory has demonstrated (past tense) that, at one point in time, they were able to follow an approved procedure and to obtain results for a specific analyte to within a predefined margin of error. The laboratory is consequently certified to generate data with respect to that specific compound using a particular method. Certification necessarily demands that the laboratory, and what it does, be placed in a black box, fundamentally removing it from the investigation process, and consequently certification is ill-equipped to convey a judgment with respect to data usability (which are inherently site-specific).

ECONOMIC DRIVERS

Both data generators and data users are under economic pressures to drive down the cost of their respective services. As a result, data generators may take shortcuts, and data

users may circumvent the DQO process. This combination of factors is very dangerous and has lead to an untold number of situations where the end users absolute confidence in environmental data may be unwarranted. Ironically, because these short cuts can so dramatically impact price, both generators and users are rewarded by receiving additional work. This vicious cycle has lead to an increased frequency of the proliferation of "data time bombs".

Process Control and Automation

Environmental analysis is a rigorous science, but not a perfect one. Its defensibility is based in the methods and associated documentation. As a result of the methods involved, the complexity of the matrices, and an inherent uncertainty, one can expect a certain fraction of analytical runs (between 25% and 50%) to be dedicated to the performance of quality control. This quality control will come in the form of either scheduled QC such as calibrations, spikes and blanks, or unscheduled QC in the form of re-analyses, dilutions, and cleaning blanks.

In addition to this expected split between QC and environmental samples, there is a minimum amount of time a laboratory needs to dedicate to the review of generated data. This can be a time-consuming process which cannot and should not be automated. Red flags should be raised in the eyes of the data user when extremely low pricing structures are observed. Unfortunately, the data user is too often motivated purely by budgetary concerns and does not take the required time to examine data.

To meet the demand for these extremely low prices, methods may be modified, processes may be automated, QC runs may be reduced, or data review may be eliminated by the laboratory. As with any business, time is money, and if an analytical run can be reduced from 40 minutes to 30 minutes, there will necessarily be more money in the laboratory's hands (the situation is particularly dangerous when these types of shortcuts are required just to break even). However, the impact of the method modifications on quality is rarely expressed, if they are mentioned to the data user at all. In any case, the overall quality of the data will necessarily be impacted, and without a DQO process, completed with verification and validation, the data generated by these substandard analyses will peacefully exist in site investigation, remedial investigation, remedial design, and risk assessment reports until someone takes the time to look beyond the numbers on the page to how the data were generated by which time it is too late.

Personnel

In no other way can a laboratory impact the delivered cost of data than by either reducing either the number of personnel or the experience of the personnel. As important as quality control is to monitor an analytical process, it is truly less than half the quality picture. Experienced personnel are necessary to successfully carry out procedures and interpret analytical results. For every procedure which is monitored by a quality control procedure, there are two which are unmonitored. Many of these procedures fall into general laboratory procedures and good laboratory practice, but they are critically important to accurate and precise analytical results. For example, procedures such as sample transfers from one container to another, use of a syringe, or performing a dilution can impact sample results by orders of magnitude if not performed correctly.

For more complex analyses, chemists are required to make interpretations of chromatograms and spectra that can drastically impact sample results. During pesticide/PCB analysis, for example, the analyst has to have the experience to be able to differentiate PCBs from each other as well as chlordane and toxaphene (all multi-component mixtures). This differentiation cannot, and should not be made by an entry-level technician. In short, laboratory automation and QC procedures are not substitutes for experienced personnel, and experienced personnel draw premium wages. For this reason, there is a certain economic limit to successful generation of quality data, below which the data should be regarded as

suspect.

DATA TIME BOMBS

The problems described in the introduction of this paper, namely "data time bombs", occur because the DQO process outlined above does not occur. Environmental data go from the instrument to a data table where they are incorporated into reports and conclusions are drawn. Very often the DQO process is totally circumvented. The following sections describe several not so uncommon problems which can lead to data non-usability. All the examples are based on documentable real life experiences; however, they have been generalized to protect the guilty. It should also be noted that this discussion focuses on "laboratory" and "method" issues. It does not go into the issue of getting the sample to the laboratory. Sample collection is extremely important when determining the viability of the final numbers. If a sample was improperly collected, it doesn't matter how much QC the laboratory performs, the sample results are still suspect. Very few QC procedures, however, exist which evaluate sampling procedures, and as a result these problems are often not identified.

Method Selection

Problems with respect to method selection fall into two categories: conscious and unconscious. Conscious being that data user who requests an inappropriate method for a given data use. In this circumstance the data user is generally not aware of the limitation(s) of a given analytical method and uses the data for an inappropriate purpose. Unconscious method selection problems occur when a data user requests a specific method and that method is modified in a manner which negatively impacts the usability of the data, and that modification is not brought to the attention of the data user. Improper method selection is perhaps the most dangerous of "data time bomb" scenarios because the data, as a result of verification, may be of good quality, but as a result of validation may be deemed unusable.

Conscious. The problem of "conscious" improper method selection arises when there are multiple method choices to analyze for a single compound or group of compounds (which is most often the case), and the data user does not understand the strengths and limitations of each. This problem has been illustrated in the past with respect to a pentachlorophenol (PCP) contaminated site. The data user, in order to save money, chose USEPA Method 420 (Total Phenolics) to generate data with respect to PCP. For quarter after quarter the data user submitted "non-detect" data to its client assuming there was no problem at the site. The data user, however, was unaware that method 420 systematically produces false negatives with respect to saturated phenols (PCP is a saturated phenol). Method 420 was performed by the book by a competent laboratory, and all the QC criteria were within limits, yet in the final analysis, the data were unusable. The data user was ready to present several years of data in court and to argue that the site was clean. Days before the court data however, an experienced data reviewer pointed out the inappropriateness of the method, and that USEPA Method 525 (GC/MS) would have been more suitable selection.

Another common example of this problem is the selection of USEPA method 624 for the analysis of chlorinated solvents in groundwater. The practical quantitation limit for this method is 5 - 10 μg/L depending on the compound, but the Federal Maximum Contamination Level (MCL) for vinyl chloride is 2 μg/L. In order to perform an adequate contamination assessment or risk evaluation USEPA Method 524.1 is required to reach the concentrations of interest.

Unconscious. Unconscious improper method selection problems arise when the data user selects an appropriate method, but the data are compromised at the laboratory by method modification or shortcut. An example of this type of problem came to light when a fixed-base laboratory used an unapproved sonication method to extract PCBs from soils over an approved 24-hour soxlet extraction. Like the situation above, the modification systematically produced false negatives while QC data indicated the system was in control.

Ironically, the only reason the problem was exposed was because a field laboratory was present at the site and indicated the presence of PCBs at high concentrations. Because of the stigma associated with field laboratories, however, the immediate reaction was the fixed-base laboratory must be right and the field laboratory must be wrong. Only after intense scrutiny of both laboratories was the field data accepted over the fixed base data.

Method Detection Limits, Sample Quantitation Limits, Action Level and ARARs

Perhaps the fastest way data can be deemed unusable is if a sample quantitation limit exceeds the action level or ARAR for a given compound as illustrated above in the improper method selection section. Unfortunately, this problem is not always predictable ahead of time by a simple comparison of the method quantitation limit with the action level or ARAR. Many standard procedures in a laboratory will elevate a sample quantitation limit from the method detection limit stated in the literature. These procedures include dilutions, re-analyses due to matrix interferences, and adjustments for percent solids, and are completely legitimate. This last example recently impacted a remediation project in New England when Total Petroleum Hydrocarbons (TPH) were analyzed on-site by an infrared method. Because the practical quantitation for the method was 40 mg/kg and the Action Limit was 50 mg/kg the project team assumed the results would be acceptable. Soils on the site were later determined to contain 25% - 40% moisture thereby elevating the sample quantitation limits to 50 - 66 mg/kg. As a result the method selected was ill-equipped to make the determination of contaminated versus clean soils and deemed unusable.

Similar problems may arise in complex matrices such as transformer oils or samples containing multiple contaminants at disparate concentrations. Analytical methods are often written to allow a laboratory to dilute a sample in order to bring the analyte of highest concentration into the dynamic range of the instrument while simultaneously diluting concentrations with respect to all other analytes. The analyte which drives the dilution does not even have to be a target analyte. This raises the possibility that all target analytes could be presented as non-detected at a concentration well above an ARAR or action level. This situation often occurs at sites containing both petroleum and solvent contamination. Samples will be diluted for the presence of hydrocarbons (even a non-target hydrocarbon), but the site decision criteria are based on the presence of chlorinated compounds such as tetrachloroethene, trichloroethene, or vinyl chloride.

Artifacts (Chasing Ghosts)

Chasing ghosts occurs when investigations are conducted around the presence of laboratory or field artifacts. It is not uncommon for an entire risk assessments to be designed around the presence of acetone or methylene chloride before the presence of the compounds is attributed to method blank contamination. Method blank contamination is generally fairly obvious and simple to correct, but there are documented cases where an entire remedial investigation was conducted based on the presence of an introduced compound. Several years ago during concurrent groundwater investigations on Cape Cod and Long Island, methyl ethyl ketone (MEK) began showing up in seemingly unrelated patterns across independent sites. Tens of thousands of dollars had been spent to investigate the presence of the solvent when an observant chemist noticed certain tentatively identified compounds in the mass spectra of the groundwater samples containing the MEK. Ultimately, the presence of MEK was linked to the use of methyl hydrate, a decontamination fluid containing ethanol and methanol.[9]

SUMMARY

The review of environmental data must be regarded as a risk management tool. The success of any environmental investigation is dependent, although not solely, on the chemical data, and these data are often negatively influenced by economic pressures or a

misunderstanding of the methods involved and the type of data they produce. Millions of dollars are riding on these data which are often given labels of "good" or "bad" based merely on an assumption of data quality rather than a rigorous data review process. In reality, the validity and/or usability of data have little to do with the preconceived notions that certified laboratories generate good data, or that field laboratories only generate screening data. Further, the mere performance of quality control procedure does not insure valid or usable data. Quality control should be regarded as a means to an end where data validity and usability can truly be evaluated. Questions of data quality, data validity, data usability, and data defensibility can only be answered by using and embracing the DQO process as a necessary tool, rather than a budgetary constraint. Unfortunately, all data are not created equally and the DQO process is the only certain means by which the data user is certain to generate data of sufficient quality to meet an intended purpose. Most importantly, the DQO process is value-neutral with respect to where data are generated. It is only concerned with how data are generated and for what purpose.

The DQO process does not need to be regarded as a complex, cumbersome, budget draining undertaking. If fact it starts by asking the simple question "how do you intend to use the data?" Unfortunately, this single question is often not asked, and as a result "data time bombs" are planted into reports. By asking this first question and working with a knowledgeable fixed or field laboratory, many data problems can be averted. Data generation cannot be viewed as a "black box" where samples go in and numbers come out with little regard as to how they were generated. Data users must be compelled to participate in the data generation process regardless of where those analyses occur. By doing so, the quality (i.e., usability) of the analytical data will necessarily will improve. With this approach there are "no surprises" and no "data time bombs" waiting in the wings weeks, months, or years down the road.

REFERENCES

1. *Guidance for the Data Quality Objectives Process.* USEPA Document EPA QA/G-4. Washington, DC. September, 1994.

2. *Data Quality Objectives Process for Superfund.* USEPA Document EPA 540/R-93-071. Washington, DC. September 1993.

3. *Requirement for Quality Control of Analytical Data.* HAZWRAP. DOE/HWP-65-R1. Oak Ridge, TN. July, 1990.

4. *Guidance For Conducting Remedial Investigations and Feasibility Studies Under CERCLA.* USEPA Document EPA/540/G-89/004. Washington, DC. October, 1988.

5. *Data Quality Objectives for Remedial Response Activities Development Process.* USEPA Document EPA/540/G-87/003. Washington, DC. March, 1987.

6. National *Functional Guidelines for Evaluating Organic Analyses*, USEPA, December, 1990.

7. *Functional Guidelines for Evaluating Inorganic Analyses*", USEPA, March, 1987.

8. *Guidance for the Data Quality Objectives Process.* USEPA Document EPA QA/G-4. Washington, DC. September, 1994.

9. *Evaluation of Methyl Ethyl Ketone in Groundwater Samples.* IRP Program. Suffolk Air National Guard. HAZWRAP. Oak Ridge, TN. March, 1990.